1. a. For regulation parameter c, I picked c = 100, 150, 200, 250, 300 and 350.
   For c = 100

```
confusion matrix
[[37  3]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.93      0.94        40
           2       0.94      0.96      0.95        50

    accuracy                           0.94        90
   macro avg       0.94      0.94      0.94        90
weighted avg       0.94      0.94      0.94        90
```

c = 150

```
confusion matrix
[[37  3]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.93      0.94        40
           2       0.94      0.96      0.95        50

    accuracy                           0.94        90
   macro avg       0.94      0.94      0.94        90
weighted avg       0.94      0.94      0.94        90
```

c = 200

```
confusion matrix
[[37  3]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.93      0.94        40
           2       0.94      0.96      0.95        50

    accuracy                           0.94        90
   macro avg       0.94      0.94      0.94        90
weighted avg       0.94      0.94      0.94        90
```

c = 250

```
confusion matrix
[[37  3]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.93      0.94        40
           2       0.94      0.96      0.95        50

    accuracy                           0.94        90
   macro avg       0.94      0.94      0.94        90
weighted avg       0.94      0.94      0.94        90
```

c = 300

```
confusion matrix
[[38  2]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.95      0.95        40
           2       0.96      0.96      0.96        50

    accuracy                           0.96        90
   macro avg       0.95      0.95      0.96        90
weighted avg       0.96      0.96      0.96        90
```
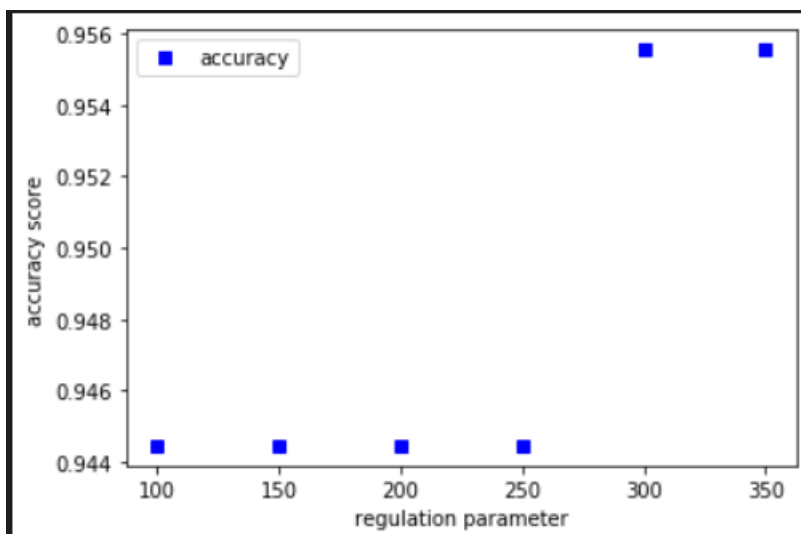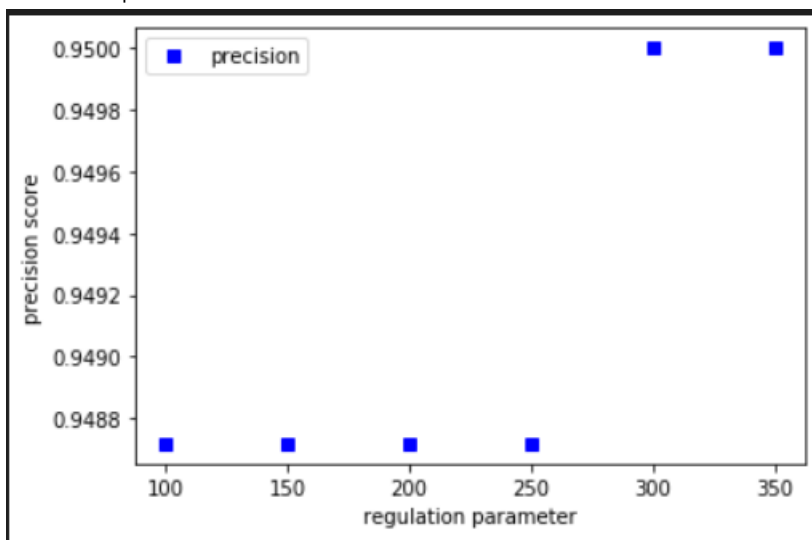
c = 350

```
confusion matrix
[[38  2]
 [ 2 48]]

classification report
              precision    recall  f1-score   support

           1       0.95      0.95      0.95        40
           2       0.96      0.96      0.96        50

    accuracy                           0.96        90
   macro avg       0.95      0.95      0.96        90
weighted avg       0.96      0.96      0.96        90
```
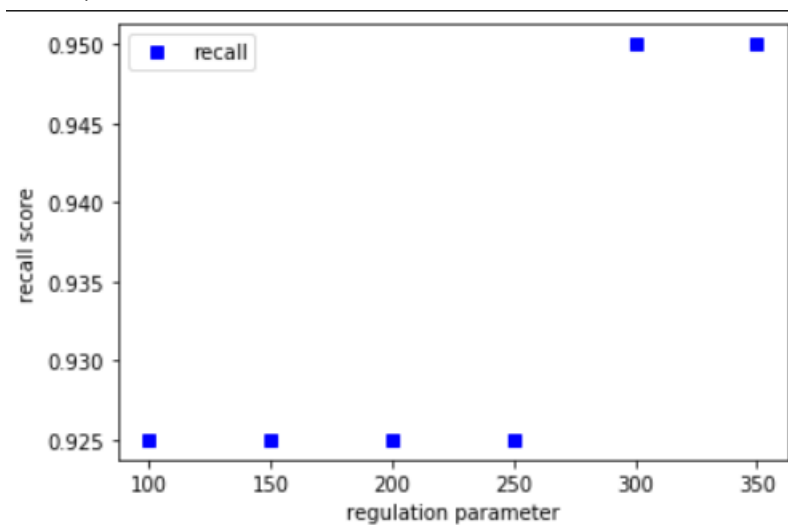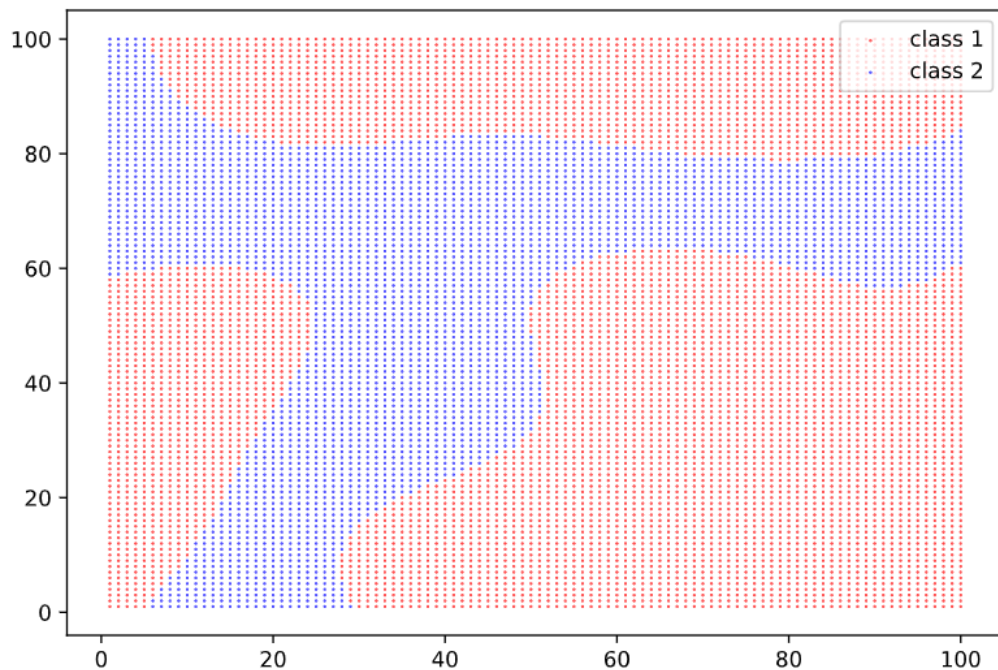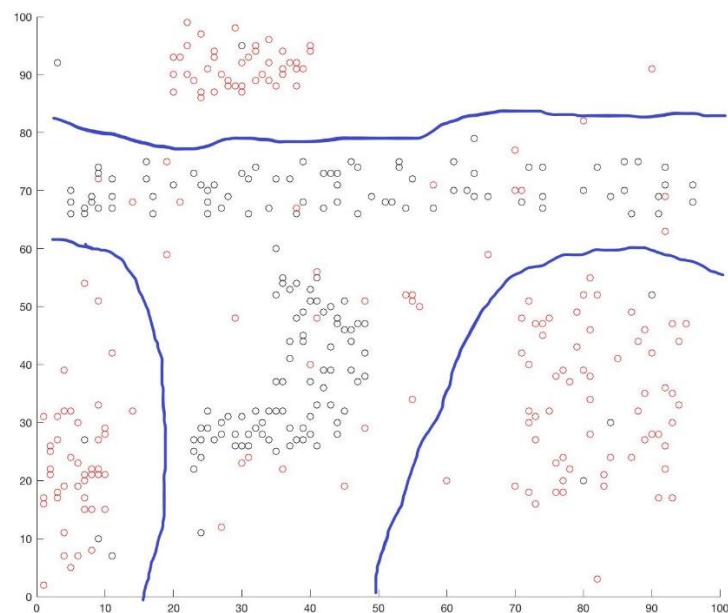
Accuracy plot

Precision plot

Recall plot

b. Regulation parameter c = 300 was chosen to be the best case.

c. i. These boundaries generally caught the t-shaped distribution, and the number of misclassifications is low. Therefore, these boundaries are efficient and effective.

ii. Same data type should be close to each other, and the margin should be as large as possible.

iii. SVM. The original data has clear boundaries in between clusters. The SVM algorithm can maximize the margin along the boundaries.

d. If the sample size if fixed, the regulation parameter c sets the amount of penalty for misclassification. A larger c will lead to a harder/smaller margin and fewer misclassification, while a smaller c will lead to a softer/larger margin and more misclassification.